

Research Article

Open Access

George Hutchinson*

On the cardinality of complex matrix scalings

DOI 10.1515/spma-2016-0014

Received November 13, 2015; accepted January 30, 2016

Abstract: We disprove a conjecture made by Rajesh Pereira and Joanna Boneng regarding the upper bound on the number of doubly quasi-stochastic scalings of an $n \times n$ positive definite matrix. In doing so, we arrive at the true upper bound for 3×3 real matrices, and demonstrate that there is no such bound when $n \geq 4$.

Keywords: Diagonal Matrix Scalings, Positive Definite Matrices, Circulant Matrices, Doubly Stochastic Matrices

MSC: 15B48, 15B51, 81P40

1 Notation and Preliminary Results

We begin by introducing some terminology. A matrix is said to be positive definite if it is Hermitian with positive eigenvalues.

Definition 1.1. Given a positive definite $n \times n$ matrix A , we say that A is doubly quasi-stochastic (DQS) if the entries in any given row (or column) sum to 1.

If A is a DQS matrix and all entries of A are real and non-negative, we say that A is doubly stochastic. (Indeed, doubly quasi-stochastic matrices are sometimes called generalized doubly stochastic matrices.) Our paper concerns the idea of doubly quasi-stochastic *scalings*. Given a positive definite matrix A and an invertible diagonal matrix D , we say that D scales A if $B = D^*AD$ is a doubly quasi-stochastic matrix. The set of all such scalings B will be denoted $sc(A)$, ie.

$$sc(A) = \{B : B \text{ is a doubly quasi-stochastic matrix and } B = D^*AD \text{ for some diagonal matrix } D\}.$$

The idea of real (positive) matrix scalings were first introduced by Sinkhorn in [1], and they have been extensively studied for their applications to a variety of topics, such as Markov processes. (See [1], [2], [3], [4] for notable work on positive scalings, or [5] for a more recent treatment of the subject.)

In [6], the authors showed that if we extend the idea to allow *complex* matrix scalings, we obtain a new application to the study of quantum entanglement. They showed that the set $sc(A)$ can be used to calculate the geometric measure of entanglement for a symmetric state whose Gram matrix is A . To this end, the authors of [6] were interested in how many matrices must be considered (ie. the cardinality of $sc(A)$). In particular, the following conjecture was made:

Conjecture 1.2. Given a positive definite $n \times n$ matrix A , $|sc(A)| \leq 2^{n-1}$.

As shown in [6], this conjecture holds true for 2×2 matrices and some special cases of higher dimensions (eg. real $n \times n$ tridiagonal matrices).

*Corresponding Author: George Hutchinson: Department of Mathematics & Statistics, University of Guelph, Guelph, Ontario, Canada N1G 2W1. E-mail: ghutch02@uoguelph.ca

In this paper, we will show that Conjecture 1.2 does not hold in general for $n \geq 3$. In Section 2, we arrive at the true upper bound for 3×3 real matrices, and give conditions on when this upper bound is attained. Section 3 is devoted to showing that there is no such upper bound when $n \geq 4$. This is achieved by giving an example of an $n \times n$ positive definite matrix with infinitely many scalings.

We take this opportunity to remind the reader of the concept of a *circulant matrix*.

Definition 1.3. An $n \times n$ matrix $A = (a_{ij})$ is called *circulant* if $a_{i_1 j_1} = a_{i_2 j_2}$ whenever $j_1 - i_1 = j_2 - i_2$ where subtraction is taken modulo n .

It is well-known (see, for example [7]) that the eigenvectors of the circulant matrix are $v_i = \frac{1}{\sqrt{n}}(1, \omega_i, \omega_i^2, \dots, \omega_i^{n-1})$ where $\{\omega_i\}_{i=1}^n$ is the set of n -th roots of unity. As shown in [6], we can use the fact that $D_i = \sqrt{n}(\text{diag}(v_i))$ is unitary to show that if A is a DQS circulant matrix with no zero entries, each of these D_i will correspond to a scaling of A . That is, $E_i^* A E_i$ is a unique scaling of A , where $E_i = k D_i$, k an appropriately chosen real constant. This provides us with n elements of $sc(A)$. We will show that in the 3×3 real case, there is almost always 3 more (that is, $sc(A) = 6$).

The remainder of this section will introduce some preliminary results that will be referenced throughout the sections that follow.

Firstly, as we will be interested in the cardinality of $sc(A)$, it will help to note that given $B \in sc(A)$, our choice of D which scales A to B is not unique. Indeed, the following lemma will be of great use to us:

Lemma 1.4 ([6] Proposition 2.1). Let $A \in M_n(\mathbb{C})$ have no zero entries and let D_1, D_2 be two invertible diagonal matrices. Then $D_1^* A D_1 = D_2^* A D_2$ if and only if $D_1 = \omega D_2$, for some $\omega \in \mathbb{C}$ with $|\omega| = 1$.

We will call D_1 and D_2 equivalent if $D_1 = \omega D_2$ (for some $|\omega|^2 = 1$).

We conclude this section by introducing the following two results that will make our lives significantly easier. (Proposition 1.6 is intrinsically proven in [6], although never explicitly stated. We include a proof here for completeness).

Proposition 1.5 ([6], Proposition 2.4). Let A be a positive definite matrix. Then there is at most one positive diagonal matrix D such that $D^* A D$ is DQS.

Proposition 1.6. Given an $n \times n$ positive definite matrix A , there is (up to equivalence) at most 2^{n-1} real diagonal matrices D such that $D^* A D$ is a doubly quasi-stochastic matrix.

Proof. We will show that given an n -dimensional sign pattern $(+/-, +/-, +/- \dots, +/-)$, there is at most one real diagonal matrix $D = \text{diag}(d_1, d_2, \dots, d_n)$ such that (d_1, d_2, \dots, d_n) fits that sign pattern and D scales A uniquely. Suppose there are 2 diagonal matrices D, E with the same sign pattern, both of which scale A . Observing that $(E D^{-1}) D A D (D^{-1} E) = E A E$ is DQS, we see that $D^{-1} E$ is a positive diagonal matrix that scales $D A D$. But clearly, the identity, I is a positive diagonal matrix that scales $D A D$. By Proposition 1.5, these must be the same matrix and we obtain $E D^{-1} = I$, or $E = D$, as desired. As there are 2^n unique sign patterns, this tells us that we have maximum 2^n real matrices that scale A . Lastly, as $D A D = (-D) A (-D)$, we see that D and $-D$ scale A to the same DQS matrix, and hence we obtain at most $2^n / 2 = 2^{n-1}$ unique scalings, as desired. \square

2 Upper bound for 3x3 real matrices

In this section, we will show that Conjecture 1.2 is false for real 3x3 matrices by constructing a counterexample with 6 scalings. We will also prove that this is the true upper bound for 3x3 real positive definite matrices.

Theorem 2.1. *Let A be a real, positive definite, 3x3 matrix. Then $|sc(A)| \leq 6$.*

Let us begin by constructing a matrix with 6 scalings.

Example 2.2. *Let A be the positive definite matrix*

$$A = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.25 & 0.5 \end{pmatrix}.$$

Then one can easily verify that the following diagonal matrices scale A to 6 unique doubly quasi-stochastic matrices:

$$D_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad D_2 = 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{2\pi i/3} & 0 \\ 0 & 0 & e^{4\pi i/3} \end{pmatrix} \quad D_3 = 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{4\pi i/3} & 0 \\ 0 & 0 & e^{2\pi i/3} \end{pmatrix}$$

$$D_4 = \begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & -\frac{2\sqrt{6}}{3} & 0 \\ 0 & 0 & -\frac{2\sqrt{6}}{3} \end{pmatrix} \quad D_5 = \begin{pmatrix} -\frac{2\sqrt{6}}{3} & 0 & 0 \\ 0 & \sqrt{6} & 0 \\ 0 & 0 & -\frac{2\sqrt{6}}{3} \end{pmatrix} \quad D_6 = \begin{pmatrix} -\frac{2\sqrt{6}}{3} & 0 & 0 \\ 0 & -\frac{2\sqrt{6}}{3} & 0 \\ 0 & 0 & \sqrt{6} \end{pmatrix}$$

Remark 2.3. *It is worth mentioning that in [6], there is an incorrect result that $sc(M) \leq 4$ for 3x3 circulant M . The problem arises when, in the proof, the entries of a certain circulant $C = circ(x, y, z)$ are erroneously assumed to be real. In fact, one must choose to take either the vector $(\bar{x}y, x\bar{z}, y\bar{z})^T$ to be real, or the entry x to be real. We cannot assume both.*

We now focus on proving the upper bound in Theorem 2.1. For simplicity, we will often want to be able to assume that A has no zero entries. To this end, we deal with the zero entry case first:

Proposition 2.4. *Let $A = (a_{ij})$ be a 3x3 real positive definite doubly stochastic matrix with zero entries $a_{\alpha\beta} = a_{\beta\alpha} = 0$ for some $1 \leq \alpha \neq \beta \leq 3$. Then $|sc(A)| \leq 4$, and every matrix in $sc(A)$ is real.*

Proof. As A has (at least) 2 zero entries, $a_{\alpha\beta} = a_{\beta\alpha} = 0$, it is permutationally similar to a tridiagonal matrix (that is, $P^{-1}AP = P^*AP$ is tridiagonal, for some permutation matrix P). Now, suppose that D scales the tridiagonal matrix P^*AP . Then D^*P^*APD is doubly stochastic, and $PD^*P^*APDP^*$ has the same rows and columns as D^*P^*APD (although they will be permuted). Thus, PDP^* scales A if and only if D scales the tridiagonal matrix P^*AP . Hence they have the same number of scalings. By [[6], Proposition 3.1], a tridiagonal matrix has at most 2^{n-1} scalings, all real. Our result follows. □

We may now proceed with our proof of Theorem 2.1. For the convenience of the reader, we write out the following matrices:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \quad D^*AD = \begin{pmatrix} a_{11}|d_1|^2 & a_{12}\bar{d}_1d_2 & a_{13}\bar{d}_1d_3 \\ a_{12}d_1\bar{d}_2 & a_{22}|d_2|^2 & a_{23}\bar{d}_2d_3 \\ a_{13}d_1\bar{d}_3 & a_{23}d_2\bar{d}_3 & a_{33}|d_3|^2 \end{pmatrix}$$

(These will be useful for reference while working through the results to follow.)

The rest of our proof of Theorem 2.1 will be broken into two parts. One in which we assume that D does not satisfy a certain property (Corollary 2.6) and one in which we assume that it does (Proposition 2.7). We will need the following easy fact:

Observation: Let z_1, z_2 be complex numbers satisfying $z_1 z_2 \in \mathbb{R}$. Then either $z_2 = k \bar{z}_1$ for some real k , or $z_1 = 0$.

Proposition 2.5. Let $A = (a_{ij})$ be a 3×3 real positive definite matrix with no zero entries, and suppose $D = \text{diag}(d_1, d_2, d_3) \in \mathbb{M}_n(\mathbb{C})$ is an invertible diagonal matrix such that D^*AD is doubly quasi-stochastic. If $\frac{d_1}{a_{23}} + \frac{d_2}{a_{13}} + \frac{d_3}{a_{12}} \neq 0$, then (up to equivalence) D is real.

Proof. Assume by a suitable rotation (apply Lemma 1.4) that d_1 is positive. Let $(D^*AD)_i$ denote the i_{th} column sum of D^*AD (that is, the i_{th} entry of eD^*AD , where e is the all ones row vector $(1, \dots, 1)$). We consider the first column of the DQS matrix D^*AD :

$$(D^*AD)_1 = a_{11}|d_1|^2 + a_{12}d_1\bar{d}_2 + a_{13}d_1\bar{d}_3 = 1$$

The first term is of course real, and this means that $d_1(a_{12}\bar{d}_2 + a_{13}\bar{d}_3)$ must be real as well. By the Observation above, this means that either $a_{12}d_2 + a_{13}d_3 = 0$ or $d_1 = k(a_{12}d_2 + a_{13}d_3)$ for some real k . We will consider each of these cases separately.

Case 1: Suppose $a_{12}d_2 + a_{13}d_3 = 0$. Then $d_2 = \frac{-a_{13}d_3}{a_{12}}$. Now consider the second column of our matrix:

$$\begin{aligned} 1 &= (D^*AD)_2 \\ &= a_{12}d_1d_2 + a_{22}|d_2|^2 + a_{23}d_2\bar{d}_3 \\ &= a_{22}|d_2|^2 + d_2(a_{12}d_1 + a_{23}\bar{d}_3) \\ &= a_{22}|d_2|^2 - a_{13}d_1d_3 - \frac{a_{23}a_{13}|d_3|^2}{a_{12}} \\ &= a_{22}|d_2|^2 - \frac{a_{23}a_{13}|d_3|^2}{a_{12}} - a_{13}d_1d_3 \end{aligned}$$

whence we see that $a_{13}d_1d_3$ must be real. As a_{13} and d_1 are nonzero real numbers, we know that d_3 must be real as well. Lastly, as $d_2 = \frac{-a_{13}d_3}{a_{12}}$, d_2 must also be real. Hence D is a real matrix, as desired.

Case 2: Now suppose $a_{12}d_2 + a_{13}d_3 \neq 0$, so that

$$d_1 = k(a_{12}d_2 + a_{13}d_3) = k(a_{12}\bar{d}_2 + a_{13}\bar{d}_3), \quad (*)$$

for some real (nonzero) k (where the second equality comes from the fact that d_1 is real). Again, we consider the second column:

$$\begin{aligned} 1 &= (D^*AD)_2 \\ &= a_{12}d_1d_2 + a_{22}|d_2|^2 + a_{23}d_2\bar{d}_3 \\ &= a_{12}k(a_{12}\bar{d}_2 + a_{13}\bar{d}_3)d_2 + a_{22}|d_2|^2 + a_{23}d_2\bar{d}_3 \\ &= a_{22}|d_2|^2 + (a_{12})^2k|d_2|^2 + (ka_{12}a_{13} + a_{23})d_2\bar{d}_3 \end{aligned}$$

from which we see that $(ka_{12}a_{13} + a_{23})d_2\bar{d}_3 \in \mathbb{R}$. Note that this cannot be zero, as $ka_{13}a_{12} + a_{23} = 0$ would mean $k = -\frac{a_{23}}{a_{13}a_{12}}$. This, combining this with $(*)$, gives us $d_2 = -\frac{a_{13}}{a_{23}}d_1 - \frac{a_{13}}{a_{12}}d_3$, which contradicts our original assumption on D that $\frac{d_1}{a_{23}} + \frac{d_2}{a_{13}} + \frac{d_3}{a_{12}} \neq 0$.

Hence, as this is nonzero, we have $d_2 = ld_3$, for some $l \in \mathbb{R} \setminus \{0\}$. Combined with $(*)$, we get $d_1 = k(a_{12}l + a_{13})d_3$. This cannot be zero, as this would mean d_1 is zero, which clearly cannot be the case if D is to scale A . Thus, d_3 must be real and, in turn, d_2 is also real as desired. \square

Proposition 1.6 combined with Proposition 2.5 gives us the following.

Corollary 2.6. *Let $A = (a_{ij})$ be a 3×3 real positive definite matrix with no zero entries. Up to equivalence, there are at most 4 complex diagonal matrices $D = \text{diag}(d_1, d_2, d_3)$ which satisfy $\frac{d_1}{a_{23}} + \frac{d_2}{a_{13}} + \frac{d_3}{a_{12}} \neq 0$ and have the property that D^*AD are doubly quasi-stochastic. (These diagonal matrices may all be taken to be real).*

We now consider scalings that do satisfy this condition.

Proposition 2.7. *Let $A = (a_{ij})$ be a 3×3 real positive definite matrix with no zero entries. Up to equivalence, there are at most 2 diagonal matrices $D = \text{diag}(d_1, d_2, d_3) \in \mathbb{M}_n(\mathbb{C})$ that scale A and satisfy $\frac{d_1}{a_{23}} + \frac{d_2}{a_{13}} + \frac{d_3}{a_{12}} = 0$.*

Proof. Let us rewrite our condition in terms of d_2 :

$$d_2 = -\frac{a_{13}}{a_{23}}d_1 - \frac{a_{13}}{a_{12}}d_3$$

Again let us assume that d_1 is real and positive by multiplying D by an appropriate scalar, if necessary. We consider the first column of D^*AD :

$$\begin{aligned} 1 = (D^*AD)_1 &= a_{11}|d_1|^2 + a_{12}d_1\bar{d}_2 + a_{13}d_1\bar{d}_3 \\ &= a_{11}|d_1|^2 + a_{12}d_1\left(-\frac{a_{13}}{a_{23}}\bar{d}_1 - \frac{a_{13}}{a_{12}}\bar{d}_3\right) + a_{13}\bar{d}_3d_1 \\ &= \left(a_{11} - \frac{a_{12}a_{13}}{a_{23}}\right)|d_1|^2 - a_{13}d_1\bar{d}_3 + a_{13}d_1\bar{d}_3 \\ &= \left(a_{11} - \frac{a_{12}a_{13}}{a_{23}}\right)|d_1|^2 \end{aligned}$$

Rearranging, we get

$$|d_1|^2 = \frac{1}{a_{11} - \frac{a_{12}a_{13}}{a_{23}}} = \frac{a_{23}}{a_{11}a_{23} - a_{12}a_{13}} \quad (1)$$

Hence, $|d_1|^2$ is fixed. As d_1 is real and positive, this fixes d_1 . Now consider the third column:

$$\begin{aligned} 1 = (D^*AD)_3 &= a_{13}\bar{d}_1d_3 + a_{23}\bar{d}_2d_3 + a_{33}|d_3|^2 \\ &= a_{13}\bar{d}_1d_3 + d_3a_{23}\left(-\frac{a_{13}}{a_{23}}\bar{d}_1 - \frac{a_{13}}{a_{12}}\bar{d}_3\right) + a_{33}|d_3|^2 \\ &= a_{13}\bar{d}_1d_3 - a_{13}\bar{d}_1d_3 - \frac{a_{13}a_{23}}{a_{12}}|d_3|^2 + a_{33}|d_3|^2 \\ &= \left(a_{33} - \frac{a_{13}a_{23}}{a_{12}}\right)|d_3|^2 \end{aligned}$$

Rearranging, we get

$$|d_3|^2 = \frac{1}{a_{33} - \frac{a_{13}a_{23}}{a_{12}}} = \frac{a_{12}}{a_{33}a_{12} - a_{13}a_{23}} \quad (3)$$

and we have that $|d_3|^2$ is fixed as well.

Lastly, we look at the second column:

$$\begin{aligned} 1 = (D^*AD)_2 &= a_{12}\bar{d}_1d_2 + a_{22}|d_2|^2 + a_{23}\bar{d}_3d_2 \\ &= -\frac{a_{12}a_{13}}{a_{23}}|d_1|^2 - a_{13}d_1\bar{d}_3 + a_{22}|d_2|^2 - a_{13}d_1\bar{d}_3 - \frac{a_{23}a_{13}}{a_{12}}|d_3|^2 \\ &= 1 - a_{11}|d_1|^2 - 2a_{13}d_1\text{Re}(d_3) + a_{22}\left|\frac{a_{13}}{a_{23}}d_1 + \frac{a_{13}}{a_{12}}d_3\right|^2 - \frac{a_{23}a_{13}}{a_{12}}|d_3|^2 \end{aligned}$$

and hence

$$0 = -a_{11}|d_1|^2 - 2a_{13}d_1\operatorname{Re}(d_3) + a_{22}\left(\frac{a_{13}^2}{a_{23}^2}|d_1|^2 + 2\frac{a_{13}^2}{a_{23}a_{12}}d_1\operatorname{Re}(d_3) + \frac{a_{13}^2}{a_{12}^2}|d_3|^2\right) - \frac{a_{23}a_{13}}{a_{12}}|d_3|^2$$

Rearranging this, we get the following:

$$2a_{13}d_1\left(1 - \frac{a_{22}a_{13}}{a_{23}a_{12}}\right)\operatorname{Re}(d_3) = \left(\frac{a_{22}a_{13}^2}{a_{23}^2} - a_{11}\right)|d_1|^2 + \frac{a_{13}}{a_{12}}\left(\frac{a_{22}a_{13}}{a_{12}} - a_{23}\right)|d_3|^2 \quad (2)$$

Note that the coefficient on $\operatorname{Re}(d_3)$ is not zero. To see this, suppose the left side is zero. This means that $a_{22}a_{13} = a_{23}a_{12}$. Substituting this into the right hand side, we obtain the following:

$$\begin{aligned} 0 &= \left(\frac{a_{23}a_{12}a_{13}}{a_{23}^2} - a_{11}\right)|d_1|^2 + \frac{a_{13}}{a_{12}}\left(\frac{a_{23}a_{12}}{a_{12}} - a_{23}\right)|d_3|^2 \\ &= \left(\frac{a_{12}a_{13}}{a_{23}} - a_{11}\right)|d_1|^2 + 0 \end{aligned}$$

For this to be consistent, we need the coefficient on $|d_1|^2$ also to be zero. This means that $\frac{a_{12}a_{13}}{a_{23}} = a_{11}$. But looking back to (1), we see that this cannot be the case. Therefore the coefficient on $\operatorname{Re}(d_3)$ is not zero.

Substituting the value obtained for d_1 and $|d_3|^2$ into (2), we fix the value for $\operatorname{Re}(d_3)$. Given $|d_3|^2$ and $\operatorname{Re}(d_3)$, we have (maximum) two choices for d_3 :

$$d_3 = \operatorname{Re}(d_3) \pm \sqrt{(|d_3|^2 - (\operatorname{Re}(d_3))^2)}i.$$

d_2 is of course fixed as soon as we have d_1 and d_3 , by our original assumption on D .

Hence, we have maximum two (possibly complex) diagonal matrices D that scale A and satisfy $\frac{d_1}{a_{23}} + \frac{d_2}{a_{13}} + \frac{d_3}{a_{12}} = 0$. □

We can now easily prove Theorem 2.1:

Proof of Theorem 2.1 Suppose that A has zero entries. Then by Proposition 2.4, $|sc(A)| \leq 4$.

If A has no zero entries, then Corollary 2.6 gives us a maximum of 4 scalings that arise from diagonal matrices satisfying $\frac{d_1}{a_{23}} + \frac{d_2}{a_{13}} + \frac{d_3}{a_{12}} \neq 0$ and Proposition 2.7 gives us a maximum of 2 scalings that arise from diagonal matrices satisfying $\frac{d_1}{a_{23}} + \frac{d_2}{a_{13}} + \frac{d_3}{a_{12}} = 0$. Hence, $|sc(A)| \leq 6$. □

Corollary 2.6 and Proposition 2.7 actually give us a condition for when $sc(A)$ will contain complex matrices. Recall that for two $n \times n$ matrices A and B , we define their *Hadamard product* $C = A \circ B$ as the $n \times n$ matrix with entries $(c_{ij}) = (a_{ij}b_{ij})$,

Corollary 2.8. *A real 3×3 positive definite matrix $A = (a_{ij})$ will have all real scalings (and hence $|sc(A)| \leq 4$) if either of the following holds:*

- (1*) $C = A \circ A^{-1}$ has at least one nonnegative off-diagonal entry (ie. $c_{ij} \geq 0$ for some $i \neq j$).
- (2*)

$$1 \leq \frac{\left| \left(\frac{a_{22}a_{13}}{a_{23}^2} - \frac{a_{11}}{a_{13}} \right) \sqrt{\frac{a_{33} - \frac{a_{13}a_{23}}{a_{12}}}{a_{11} - \frac{a_{12}a_{13}}{a_{23}}} + \left(\frac{a_{22}a_{13}}{a_{12}^2} - \frac{a_{23}}{a_{12}} \right) \sqrt{\frac{a_{11} - \frac{a_{12}a_{13}}{a_{23}}}{a_{33} - \frac{a_{13}a_{23}}{a_{12}}}} \right|}{2 \left(1 - \frac{a_{22}a_{13}}{a_{23}a_{12}} \right)}$$

Otherwise $sc(A)$ will contain two complex matrices.

Proof. Let us begin by proving sufficiency of (1*). Recall that the off-diagonal entries of A^{-1} can be expressed as

$$(A^{-1})_{ij} = \frac{1}{\det(A)}(a_{ik}a_{kj} - a_{ij}a_{kk}), \quad k \neq i, j.$$

As A is positive definite (and hence has positive determinant), we can thus re-write the condition in (1^*) as:

$$a_{ij}(a_{ik}a_{kj} - a_{ij}a_{kk}) \geq 0, \text{ for all } 1 \leq i, j, k \leq 3, \text{ where } i, j, k \text{ are mutually distinct.}$$

If $a_{ij} = 0$, then Proposition 2.4 tells us that we have only real scalings. If $a_{ij} \neq 0$, we divide both sides by the positive number a_{ij}^2 to obtain

$$\frac{(a_{ik}a_{kj} - a_{ij}a_{kk})}{a_{ij}} \geq 0. \quad (*)$$

Now, suppose for the purposes of contradiction that $(*)$ holds, and A has complex scalings. From the proof of Proposition 2.7, we know that (1), (2), and (3) must hold. If, $(i, j, k) = (1, 2, 3)$ (or $(i, j, k) = (2, 1, 3)$, by symmetry), and $(*)$ combines with (3) to yield $-\frac{1}{|d_3|^2} \geq 0$. Clearly this is impossible, and we have our contradiction. Similarly, if $(i, j, k) = (2, 3, 1)$ or $(i, j, k) = (3, 2, 1)$, $(*)$ combines with (1) to yield $-\frac{1}{|d_1|^2} \geq 0$. It is easy to show that $(i, j, k) = (1, 3, 2)$ or $(i, j, k) = (3, 1, 2)$ will imply that $-\frac{1}{|d_2|^2} \geq 0$. (This can be seen by reworking the proof of Proposition 2.7, where the condition is written in terms of d_3 instead of d_2 , or simply by noting that any condition such as (1) or (3) which restricts the scalings of A must necessarily be invariant with respect to permutation-similarity.) Thus, if (1^*) holds, we cannot have complex scalings.

Let us now consider (2^*) . Again, suppose A has complex scalings, and hence (1),(2), and (3) hold.

Rearranging (2), we obtain:

$$\begin{aligned} \operatorname{Re}(d_3) &= \frac{(\frac{a_{22}a_{13}^2}{a_{23}^2} - a_{11})|d_1|^2 + \frac{a_{13}}{a_{12}}(\frac{a_{22}a_{13}}{a_{12}} - a_{23})|d_3|^2}{2a_{13}d_1(1 - \frac{a_{22}a_{13}}{a_{23}a_{12}})} \\ &= \frac{(\frac{a_{22}a_{13}}{a_{23}^2} - \frac{a_{11}}{a_{13}})d_1 + (\frac{a_{22}a_{13}}{a_{12}^2} - \frac{a_{23}}{a_{12}})\frac{|d_3|^2}{d_1}}{2(1 - \frac{a_{22}a_{13}}{a_{23}a_{12}})} \quad (**)$$

Taking the absolute value of $(**)$ and dividing by $|d_3|$ yields:

$$\left| \frac{\operatorname{Re}(d_3)}{d_3} \right| = \left| \frac{(\frac{a_{22}a_{13}}{a_{23}^2} - \frac{a_{11}}{a_{13}})\frac{d_1}{|d_3|} + (\frac{a_{22}a_{13}}{a_{12}^2} - \frac{a_{23}}{a_{12}})\frac{|d_3|}{d_1}}{2(1 - \frac{a_{22}a_{13}}{a_{23}a_{12}})} \right|.$$

Substituting our values in (1) and (3) in for d_1 and $|d_3|$ gives the right hand side of (2^*) . That is, the right hand side of (2^*) is $\left| \frac{\operatorname{Re}(d_3)}{d_3} \right|$. If d_3 is to exist and be complex, it must satisfy $\left| \frac{\operatorname{Re}(d_3)}{d_3} \right| < 1$. Thus, our complex scalings do not exist if (2^*) holds. (It is easy to see that if d_3 is real and d_1 is real, we must have that d_2 is real by examining the first row sum of D^*AD .)

If neither of our conditions hold, then we will necessarily have 2 complex scalings (obtained by simply filling in the values of (a_{ij}) into (1),(2) and (3)). □

Remark 2.9. 1) Given the relative complexity of (2^*) , one might hope that (2^*) implies (1^*) , which would allow us to remove this condition from Corollary 2.8. Unfortunately, this is not the case. Indeed, consider the following DQS matrix:

$$\begin{bmatrix} 1.9 & -1 & 0.1 \\ -1 & 1.8 & 0.2 \\ 0.1 & 0.2 & 0.7 \end{bmatrix}$$

This matrix satisfies (2^*) , and hence has only real scalings. However, it does not satisfy (1^*) .

2) Similarly, we can show that property (1*) cannot be removed from Corollary 2.8 by examining the DQS matrix:

$$A = \begin{bmatrix} 0.422 & 0.16 & 0.418 \\ 0.16 & 0.673 & 0.167 \\ 0.418 & 0.167 & 0.415 \end{bmatrix}.$$

This matrix satisfies (1*), as the (1,2)-entry of $A \circ A^{-1}$ is positive. However, it does not satisfy (2*), as the value of the right hand side of (2*) works out to be slightly less than 1 (~0.9985).

Now that we have an upper bound on the number of scalings for 2x2 positive definite matrices ($|sc(A)| \leq 2$) and 3x3 real positive definite matrices ($|sc(A)| \leq 6$), we might be tempted to modify Conjecture 1.2 and suggest a new upper bound on $|sc(A)|$ for $n \times n$ positive definite matrices. Our next section shows that such a bound does not exist for any dimension higher than $n=3$.

3 Real Matrices with infinitely many scalings

In this section, we will show that for $n \geq 4$, there exists a class of $n \times n$ real positive definite matrices C such that $|sc(C)|$ is not finite.

Theorem 3.1. *Let $n \geq 4$ and suppose $C \in \mathbb{M}_n(\mathbb{R})$ is a positive definite circulant matrix, ie.*

$$C = \text{circ}(a, b, b, b \dots b) = \begin{bmatrix} a & b & b & b & b & \dots & b \\ b & a & b & b & b & \dots & b \\ & & & & & \ddots & \\ & & & & & & \\ & & & & & & \\ b & b & b & b & \dots & b & a \end{bmatrix}$$

satisfying $a > 0$, $\frac{-a}{n-1} < b < a$, where $b \neq 0$. Then $|sc(C)|$ is not finite.

Proof. For any complex number z of modulus 1, define E_z to be the following diagonal matrix:

$$E_z = \begin{bmatrix} z & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & -z & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \omega & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \omega^2 & 0 & \dots & \\ 0 & 0 & 0 & 0 & \ddots & & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & \omega^{n-2} \end{bmatrix}$$

where ω is the $(n - 2)$ th root of unity. (ie. $\omega = e^{\frac{2\pi i}{n-2}}$). Next, let $D_z = \frac{1}{\sqrt{a-b}} E_z$ (this is of course well-defined, as C is positive definite and hence $a > b$). We claim that D_z scales C .

Indeed, $(D_z^* C D_z) = \frac{1}{a-b} (E_z^* C E_z)$, so we need only show that $E_z^* C E_z$ has row sums equal to $a - b$.

$$E_z^* C E_z = \begin{bmatrix} |z|^2 a & -|z|^2 b & \bar{z}\omega b & \bar{z}\omega^2 b & \dots & \bar{z}\omega^{n-2} b \\ -|z|^2 b & |z|^2 a & -\bar{z}\omega b & \bar{z}\omega^2 b & \dots & -\bar{z}\omega^{n-2} b \\ \bar{\omega} z b & -\bar{\omega} z b & |\omega|^2 a & \bar{\omega}\omega^2 b & \dots & \bar{\omega}\omega^{n-2} b \\ & & & \ddots & & \\ \overline{\omega^{n-2} z} b & -\overline{\omega^{n-2} z} b & \overline{\omega^{n-2} \omega} b & \dots & \overline{\omega^{n-2} \omega^{n-3}} b & |\omega^{n-2}|^2 a \end{bmatrix}$$

Now it is simply a matter of investigating each row sum.

$$(E_z^*CE_z)_1 = |z|^2a - |z|^2b + \bar{z}b(\omega + \omega^2 + \dots + \omega^{n-2}).$$

Of course, the sum of all (n-2)th roots of unity is zero, and $|z|^2 = 1$, by assumption on z . Hence this row sum is indeed $a - b$.

The terms in the second row are exactly the same as the first, so we inspect the third row next:

$$(E_z^*CE_z)_3 = \bar{\omega}(zb - zb) + |\omega|^2a + \bar{\omega}b(\omega^2 + \dots + \omega^{n-2}) = a + \bar{\omega}b(0 - \omega) = a - b$$

where we have again used the fact that the roots of unity sum to 0, and $|\omega|^2 = 1$.

It is easy to see that the other rows (rows 4 through n) will all sum to $a - b$, by the exact same reasoning:

$$(E_z^*CE_z)_r = \overline{\omega^{r-2}}(zb - zb) + |\omega^{r-2}|^2a + b\overline{\omega^{r-2}}\left(\sum_{\substack{m=1 \\ m \neq r-2}}^{n-2} \omega^m\right) = 0 + a + b\overline{\omega^{r-2}}(0 - \omega^{r-2}) = a - b.$$

Hence, D_z scales C to a DQS matrix, for all complex z of modulus 1. Further, different choices of z will necessarily give rise to different scalings (for example, the (n,1) entry of $D_z^*CD_z$ is $\left(\frac{\overline{\omega^{n-2}}zb}{a-b}\right)$, and $b \neq 0$, by assumption). Thus $sc(C)$ contains infinitely many unique scalings, corresponding to each choice of z . □

As mentioned in the introduction, the interest in complex scalings arose due in part to a connection with the geometric measure of entanglement. In particular, once a scaling was obtained, one wanted to investigate that scaling’s permanent. In this context, it is worth noting that while we have constructed infinite scalings of C above, each of these scalings will all have the same permanent (indeed, they will all have permanent $\frac{per(C)}{a-b}$).

4 Additional comments

In [2], Marshall and Olkin show that for any real positive definite matrix A , there exists a unique positive real D that scales A (see [8] for a very nice constructive proof of this result, which yields an easily implemented algorithm for finding real scalings). Note that for any diagonal matrix E consisting of 1s and -1s, EAE is still positive definite, and hence will be scaled by a positive D_E . This means that for each diagonal sign pattern of E , D_EE is a real diagonal matrix that scales A . Once one accounts for double counting (which will arise from the property described in Lemma 1.4), this shows that a real $n \times n$ matrix with no zero entries will have exactly 2^{n-1} real matrices in $sc(A)$, each with a different sign pattern.

This (combined with the fact that a particular scalar multiple of $D = \text{diag}(1, \omega, \omega^2)$ and $E = \text{diag}(1, \omega^2, \omega)$ will always scale a DQS circulant) shows that there is nothing special about the DQS circulant matrix chosen as our counterexample at the beginning of Section 2. Indeed, most (those with all nonzero entries) real DQS circulant 3x3 matrices C have $|sc(C)| = 6$. In reference to Example 2.2, a real DQS circulant 3x3 matrix with non-zero entries will be scaled by D_1 , scalar multiples of D_2 and D_3 , one diagonal matrix with the sign pattern of D_4 , and two matrices which are simple rotations of this (as D_5 and D_6 are rotations of D_4).

Acknowledgement: The author wishes to sincerely thank Dr. Rajesh Pereira for many fruitful discussions and suggestions on how to improve this paper, as well as the anonymous referees for their helpful feedback and constructive comments.

References

- [1] R. Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. *Annals of Mathematical Statistics*, 35(2):876–879, 1964.
- [2] A. W. Marshall and I. Olkin. Scaling of matrices to achieve specified row and column sums. *Numer. Math.*, 12(1):83–90, 1968.
- [3] M. V. Menon. Reduction of a matrix with positive elements to a doubly stochastic matrix. *Proc. Amer. Math. Soc.*, 18:244–247, 1967.
- [4] R. Brualdi, S. Parter, and H. Schneider. The diagonal equivalence of a non-negative matrix to a stochastic matrix. *J. Math. Anal. Appl.*, 16:31–50, 1966.
- [5] C. R. Johnson and R. Reams. Scaling of symmetric matrices by positive diagonal congruence. *Linear Multilinear Algebra*, 57:123–140, 2009.
- [6] R. Pereira and J. Boneng. The theory and applications of complex matrix scalings, *Spec. Matrices*, 2: 68-77, 2014
- [7] P. J. Davis. *Circulant Matrices*. John Wiley & Sons, 1979.
- [8] D. P. O’Leary. Scaling symmetric positive definite matrices to prescribed row sums. *Linear Algebra Appl.*, pages 185–191, 2003.